D. Romeo

RAW SEQUENCE LISTING   DATE: 05/20/2003
PATENT APPLICATION: **US/09/716,147**   TIME: 07:51:00

Input Set : **N:\Crf3\RULE60\09716147.RAW.txt**
Output Set: **N:\CRF4\05202003\I716147.raw**

SEQUENCE LISTING

```
 1  (1) GENERAL INFORMATION:
 2       (i) APPLICANT: Hinuma, Shuji
 3                     Habata, Yugo
 4                     Kawamata, Yuji
 5                     Hosoya, Masaki
 6                     Fujii, Ryo
 7                     Fukusumi, Shoji
 8                     Kitada, Chieko
 9      (ii) TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
10     (iii) NUMBER OF SEQUENCES: 140
11      (iv) CORRESPONDENCE ADDRESS:
12           (A) ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
13           (B) STREET: 130 Water Street
14           (C) CITY: Boston
15           (D) STATE: MA
16           (E) COUNTRY: USA
17           (F) ZIP: 02109
18       (v) COMPUTER READABLE FORM:
19           (A) MEDIUM TYPE: Diskette
20           (B) COMPUTER: IBM compatible
21           (C) OPERATING SYSTEM: DOS
22           (D) SOFTWARE: FastSEQ for Windows Version 2.0
23      (vi) CURRENT APPLICATION DATA:
C--> 24           (A) APPLICATION NUMBER: US/09/716,147
C--> 25           (B) FILING DATE: 17-Nov-2000
26           (C) CLASSIFICATION:
27     (vii) PRIOR APPLICATION DATA:
W--> 28           (A) APPLICATION NUMBER: US/08/776,971B
29           (B) FILING DATE: 06-Feb-1997
W--> 30           (A) APPLICATION NUMBER: PCT/JP96/03821
31           (B) FILING DATE: 28-DEC-1996
W--> 32           (A) APPLICATION NUMBER: JP 7/343371
33           (B) FILING DATE: 28-DEC-1995
W--> 34           (A) APPLICATION NUMBER: JP 8/59419
35           (B) FILING DATE: 15-MAR-1996
W--> 36           (A) APPLICATION NUMBER: JP 8/211805
37           (B) FILING DATE: 12-AUG-1996
W--> 38           (A) APPLICATION NUMBER: JP 8/246573
39           (B) FILING DATE: 18-SEP-1996
40    (viii) ATTORNEY/AGENT INFORMATION:
41           (A) NAME: Conlin, David G.
42           (B) REGISTRATION NUMBER: 27,026
```

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: **US/09/716,147**

DATE: 05/20/2003
TIME: 07:51:00

Input Set : **N:\Crf3\RULE60\09716147.RAW.txt**
Output Set: **N:\CRF4\05202003\I716147.raw**

```
43              (C) REFERENCE/DOCKET NUMBER: 47176
44      (ix) TELECOMMUNICATION INFORMATION:
45              (A) TELEPHONE: 617-523-3400
46              (B) TELEFAX: 617-523-6440
47  (2) INFORMATION FOR SEQ ID NO: 1:
48      (i) SEQUENCE CHARACTERISTICS:
49              (A) LENGTH: 98 amino acids
50              (B) TYPE: amino acid
51              (C) STRANDEDNESS: single
52              (D) TOPOLOGY: linear
53      (ii) MOLECULE TYPE: protein
54      (v) FRAGMENT TYPE: internal
55      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
56       Met Lys Ala Val Gly Ala Trp Leu Leu Cys Leu Leu Leu Leu Gly Leu
57       1               5                   10                  15
58       Ala Leu Gln Gly Ala Ala Ser Arg Ala His Gln His Ser Met Glu Ile
59                   20                  25                  30
60       Arg Thr Pro Asp Ile Asn Pro Ala Trp Tyr Ala Gly Arg Gly Ile Arg
61               35                  40                  45
62       Pro Val Gly Arg Phe Gly Arg Arg Arg Ala Ala Pro Gly Asp Gly Pro
63               50                  55          ·       60
64       Arg Pro Gly Pro Arg Arg Val Pro Ala Cys Phe Arg Leu Glu Gly Gly
65       65                  70                  75                  80
66       Ala Glu Pro Ser Arg Ala Leu Pro Gly Arg Leu Thr Ala Gln Leu Val
67                   85                  90                  95
68       Gln Glu
70  (2) INFORMATION FOR SEQ ID NO: 2:
71      (i) SEQUENCE CHARACTERISTICS:
72              (A) LENGTH: 294 base pairs
73              (B) TYPE: nucleic acid
74              (C) STRANDEDNESS: double
75              (D) TOPOLOGY: linear
76      (ii) MOLECULE TYPE: cDNA
77      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
78       ATGAAGGCGG TGGGGGCCTG GCTCCTCTGC CTGCTGCTGC TGGGCCTGGC CCTGCAGGGG    60
79       GCTGCCAGCA GAGCCCACCA GCACTCCATG GAGATCCGCA CCCCCGACAT CAACCCTGCC   120
80       TGGTACGCRG GCCGTGGGAT CCGGCCCGTG GGCCGCTTCG GCCGGCGAAG AGCTGCCCYG   180
81       GGGGACGGAC CCAGGCCTGG CCCCCGGCGT GTGCCGGCCT GCTTCCGCCT GGAAGGCGGY   240
82       GCTGAGCCCT CCCGAGCCCT CCCGGGGCGG CTGACGGCCC AGCTGGTCCA GGAA          294
84  (2) INFORMATION FOR SEQ ID NO: 3:
85      (i) SEQUENCE CHARACTERISTICS:
86              (A) LENGTH: 29 amino acids
87              (B) TYPE: amino acid
88              (C) STRANDEDNESS: single
89              (D) TOPOLOGY: linear
90      (ii) MOLECULE TYPE: protein
91      (v) FRAGMENT TYPE: internal
92      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
93       Ser Arg Ala His Gln His Ser Met Glu Ile Arg Thr Pro Asp Ile Asn
```

RAW SEQUENCE LISTING

PATENT APPLICATION: **US/09/716,147**

DATE: 05/20/2003
TIME: 07:51:00

Input Set : **N:\Crf3\RULE60\09716147.RAW.txt**
Output Set: **N:\CRF4\05202003\I716147.raw**

```
94        1             5                    10                       15
95        Pro Ala Trp Tyr Ala Gly Arg Gly Ile Arg Pro Val Gly
96                       20                   25
98  (2) INFORMATION FOR SEQ ID NO: 4:
99       (i) SEQUENCE CHARACTERISTICS:
100           (A) LENGTH: 19 amino acids
101           (B) TYPE: amino acid
102           (C) STRANDEDNESS: single
103           (D) TOPOLOGY: linear
104      (ii) MOLECULE TYPE: protein
105       (v) FRAGMENT TYPE: internal
106      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
107          Thr Pro Asp Ile Asn Pro Ala Trp Tyr Ala Gly Arg Gly Ile Arg Pro
108          1             5                    10                       15
109          Val Gly Arg
111  (2) INFORMATION FOR SEQ ID NO: 5:
112       (i) SEQUENCE CHARACTERISTICS:
113           (A) LENGTH: 31 amino acids
114           (B) TYPE: amino acid
115           (C) STRANDEDNESS: single
116           (D) TOPOLOGY: linear
117      (ii) MOLECULE TYPE: protein
118       (v) FRAGMENT TYPE: internal
119      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
120          Ser Arg Ala His Gln His Ser Met Glu Ile Arg Thr Pro Asp Ile Asn
121          1             5                    10                       15
122          Pro Ala Trp Tyr Ala Gly Arg Gly Ile Arg Pro Val Gly Arg Phe
123                       20                   25                       30
125  (2) INFORMATION FOR SEQ ID NO: 6:
126       (i) SEQUENCE CHARACTERISTICS:
127           (A) LENGTH: 32 amino acids
128           (B) TYPE: amino acid
129           (C) STRANDEDNESS: single
130           (D) TOPOLOGY: linear
131      (ii) MOLECULE TYPE: protein
132       (v) FRAGMENT TYPE: internal
133      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
134          Ser Arg Ala His Gln His Ser Met Glu Ile Arg Thr Pro Asp Ile Asn
135          1             5                    10                       15
136          Pro Ala Trp Tyr Ala Gly Arg Gly Ile Arg Pro Val Gly Arg Phe Gly
137                       20                   25                       30
139  (2) INFORMATION FOR SEQ ID NO: 7:
140       (i) SEQUENCE CHARACTERISTICS:
141           (A) LENGTH: 33 amino acids
142           (B) TYPE: amino acid
143           (C) STRANDEDNESS: single
144           (D) TOPOLOGY: linear
145      (ii) MOLECULE TYPE: protein
146       (v) FRAGMENT TYPE: internal
```

```
147   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
148     Ser Arg Ala His Gln His Ser Met Glu Ile Arg Thr Pro Asp Ile Asn
149     1               5                   10                  15
150     Pro Ala Trp Tyr Ala Gly Arg Gly Ile Arg Pro Val Gly Arg Phe Gly
151                 20                  25                  30
152     Arg
154 (2) INFORMATION FOR SEQ ID NO: 8:
155     (i) SEQUENCE CHARACTERISTICS:
156           (A) LENGTH: 20 amino acids
157           (B) TYPE: amino acid
158           (C) STRANDEDNESS: single
159           (D) TOPOLOGY: linear
160    (ii) MOLECULE TYPE: protein
161     (v) FRAGMENT TYPE: internal
162    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
163     Thr Pro Asp Ile Asn Pro Ala Trp Tyr Ala Gly Arg Gly Ile Arg Pro
164     1               5                   10                  15
165     Val Gly Arg Phe
166                 20
168 (2) INFORMATION FOR SEQ ID NO: 9:
169     (i) SEQUENCE CHARACTERISTICS:
170           (A) LENGTH: 21 amino acids
171           (B) TYPE: amino acid
172           (C) STRANDEDNESS: single
173           (D) TOPOLOGY: linear
174    (ii) MOLECULE TYPE: protein
175     (v) FRAGMENT TYPE: internal
176    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
177     Thr Pro Asp Ile Asn Pro Ala Trp Tyr Ala Gly Arg Gly Ile Arg Pro
178     1               5                   10                  15
179     Val Gly Arg Phe Gly
180                 20
182 (2) INFORMATION FOR SEQ ID NO: 10:
183     (i) SEQUENCE CHARACTERISTICS:
184           (A) LENGTH: 22 amino acids
185           (B) TYPE: amino acid
186           (C) STRANDEDNESS: single
187           (D) TOPOLOGY: linear
188    (ii) MOLECULE TYPE: protein
189     (v) FRAGMENT TYPE: internal
190    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
191     Thr Pro Asp Ile Asn Pro Ala Trp Tyr Ala Gly Arg Gly Ile Arg Pro
192     1               5                   10                  15
193     Val Gly Arg Phe Gly Arg
194                 20
196 (2) INFORMATION FOR SEQ ID NO: 11:
197     (i) SEQUENCE CHARACTERISTICS:
198           (A) LENGTH: 87 base pairs
199           (B) TYPE: nucleic acid
```

```
200          (C) STRANDEDNESS: double
201          (D) TOPOLOGY: linear
202     (ii) MOLECULE TYPE: cDNA
203     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:                              60
204       AGCAGAGCCC ACCAGCACTC CATGGAGATC CGCACCCCCG ACATCAACCC TGCCTGGTAC     87
205       GCRGGCCGTG GGATCCGGCC CGTGGGC
207  (2) INFORMATION FOR SEQ ID NO: 12:
208      (i) SEQUENCE CHARACTERISTICS:
209          (A) LENGTH: 57 base pairs
210          (B) TYPE: nucleic acid
211          (C) STRANDEDNESS: double
212          (D) TOPOLOGY: linear
213     (ii) MOLECULE TYPE: cDNA
214     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:                              57
215       ACCCCCGACA TCAACCCTGC CTGGTACGCR GGCCGTGGGA TCCGGCCCGT GGGCCGC
217  (2) INFORMATION FOR SEQ ID NO: 13:
218      (i) SEQUENCE CHARACTERISTICS:
219          (A) LENGTH: 93 base pairs
220          (B) TYPE: nucleic acid
221          (C) STRANDEDNESS: double
222          (D) TOPOLOGY: linear
223     (ii) MOLECULE TYPE: cDNA
224     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:                              60
225       AGCAGAGCCC ACCAGCACTC CATGGAGATC CGCACCCCCG ACATCAACCC TGCCTGGTAC     93
226       GCRGGCCGTG GGATCCGGCC CGTGGGCCGC TTC
228  (2) INFORMATION FOR SEQ ID NO: 14:
229      (i) SEQUENCE CHARACTERISTICS:
230          (A) LENGTH: 96 base pairs
231          (B) TYPE: nucleic acid
232          (C) STRANDEDNESS: double
233          (D) TOPOLOGY: linear
234     (ii) MOLECULE TYPE: cDNA
235     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:                              60
236       AGCAGAGCCC ACCAGCACTC CATGGAGATC CGCACCCCCG ACATCAACCC TGCCTGGTAC     96
237       GCRGGCCGTG GGATCCGGCC CGTGGGCCGC TTCGGC
239  (2) INFORMATION FOR SEQ ID NO: 15:
240      (i) SEQUENCE CHARACTERISTICS:
241          (A) LENGTH: 99 base pairs
242          (B) TYPE: nucleic acid
243          (C) STRANDEDNESS: double
244          (D) TOPOLOGY: linear
245     (ii) MOLECULE TYPE: cDNA
246     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:                              60
247       AGCAGAGCCC ACCAGCACTC CATGGAGATC CGCACCCCCG ACATCAACCC TGCCTGGTAC     99
248       GCRGGCCGTG GGATCCGGCC CGTGGGCCGC TTCGGCCGG
250  (2) INFORMATION FOR SEQ ID NO: 16:
251      (i) SEQUENCE CHARACTERISTICS:
252          (A) LENGTH: 60 base pairs
253          (B) TYPE: nucleic acid
```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/716,147

DATE: 05/20/2003
TIME: 07:51:01

Input Set : N:\Crf3\RULE60\09716147.RAW.txt
Output Set: N:\CRF4\05202003\I716147.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:29; N Pos. 20
Seq#:30; N Pos. 3,12
Seq#:31; N Pos. 15,16
Seq#:32; N Pos. 23
Seq#:35; N Pos. 3
Seq#:36; N Pos. 3,9,18,24
Seq#:37; N Pos. 15,21
Seq#:73; Xaa Pos.10,11,21,22
Seq#:74; Xaa Pos.3,5,10
Seq#:114; Xaa Pos.43,80
Seq#:115; Xaa Pos.43,80
Seq#:116; Xaa Pos.43,80
Seq#:117; Xaa Pos.43,80

VERIFICATION SUMMARY
PATENT APPLICATION: **US/09/716,147**
DATE: 05/20/2003
TIME: 07:51:01

Input Set : **N:\Crf3\RULE60\09716147.RAW.txt**
Output Set: **N:\CRF4\05202003\I716147.raw**

```
L:24 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:25 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:30 M:238 W: Alpha Fields not Ordered, Reordered [(A) APPLICATION NUMBER:] of (1)(vii)
L:32 M:238 W: Alpha Fields not Ordered, Reordered [(A) APPLICATION NUMBER:] of (1)(vii)
L:34 M:238 W: Alpha Fields not Ordered, Reordered [(A) APPLICATION NUMBER:] of (1)(vii)
L:36 M:238 W: Alpha Fields not Ordered, Reordered [(A) APPLICATION NUMBER:] of (1)(vii)
L:38 M:238 W: Alpha Fields not Ordered, Reordered [(A) APPLICATION NUMBER:] of (1)(vii)
L:1140 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 after pos.:0
M:341 Repeated in SeqNo=73
L:1166 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74 after pos.:0
L:1990 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:114 after pos.:146
M:341 Repeated in SeqNo=114
L:2022 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:115 after pos.:32
M:341 Repeated in SeqNo=115
L:2050 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:116 after pos.:146
M:341 Repeated in SeqNo=116
L:2082 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:117 after pos.:32
M:341 Repeated in SeqNo=117
```